DOCUMENT RESUME

ED 285 736 SE 048 324

AUTHOR	Roschelle, Jeremy; Greeno, James G.
TITLE	Mental Models in Expert Physics Reasoning.
INSTITUTION	California Univ., Berkeley.
SPONS AGENCY	Office of Naval Research, Washington, D.C. Personnel and Training Branch.
REPORT NO	GK-2
PUB DATE	Jul 87
CONTRACT	N00014-85-K-0095
NOTE	29p.; For related documents, see SE 048 323-327.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Physics; *Problem Solving; Science Education; *Science Instruction; Sciences
IDENTIFIERS	*Experts; *Mental Models

ABSTRACT

Proposed is a relational framework for characterizing experienced physicists' representations of physics problem situations and the process of constructing these representations. A representation includes a coherent set of relations among: (1) a mental model of the objects in the situation, along with their relevant properties and relations; (2) a mental model of theoretical idealizations of objects; and (3) parameter histories based on mental simulations of both models. Evidence from protocols and a small experiment support a conclusion that experienced physicists' processes of representing problem situations (a) use informal commonsense knowledge, including envisionment of objects in the situations, and (b) are interactive, with mutual influences between informal knowledge and their technical, theoretical knowledge. Also described are characteristics of the mental models that represent problem situations and the process of constructing them, drawing from work by artificial intelligence researchers on qualitative process models, and specifying several categories of rules that would be needed for an implementation of the system as a simulation program. (Author/RH)

ED285736

July 1987

Report No. GK-2

Mental Models in Expert Physics Reasoning

Jeremy Roschelle and James G. Greeno

University of California, Berkeley

This research was funded by the Office of Naval Research with Contract N00014-85-K-0095, Project NR 667-544. Approved for public release; distribution unlimited.

Mental Models in Expert Physics Reasoning

Jeremy Roschelle and James G. Greeno

University of California, Berkeley

Abstract

We propose a *relational* framework for characterizing experienced physicists' representations of physics problem situations and the process of constructing these representations. A representation includes a coherent set of relations among (1) a mental model of the objects in the situation, along with their relevant properties and relations, (2) a mental model of theoretical idealizations of objects, and (3) parameter histories based on mental simulations of both models. Evidence from protocols and a small experiment support a conclusion that experienced physicists' processes of representing problem situations (a) use informal, commonsense knowledge, including envisionment of objects in the situations, and (b) are interactive, with mutual influences between informal knowledge and their technical, theoretical knowledge. We also describe characteristics of the mental models that represent problem situations and the process of constructing them, drawing from work by AI researchers on qualitative process models, and specifying several categories of rules that would be needed for an implementation of the system as a simulation program.

This research was sponsored by the Office of Naval Research, Contract N00014-85-K-0095, Project NR-667-544. A paper based partly on these results was presented at a meeting of the Cognitive Science Society in Seattle, July 1987.

Human problem solvers often construct mental models to represent problem situations and to provide a medium for reasoning. In this paper, we discuss characteristics of mental models of situations that are described in physics problems. We report protocols of experienced physicists in tasks that were designed to emphasize the processes of representing problems. These results have led us to an understanding of some features of physics problem solving that have not been emphasized in earlier discussions.

We focus on two features of problem solving that are emphasized in the protocols. The first is the construction and use of mental models of the problem situation that provide elaborate intermediate representations. The second is the inclusion of informal, commonsense reasoning, which interacts subtly and continuously with the reasoner's formal and theoretical knowledge.

1. An Introductory Example

The diagram in Figure 1 was shown to an experienced physicist who was asked, "What's happening?" The physicist provided the following protocol:

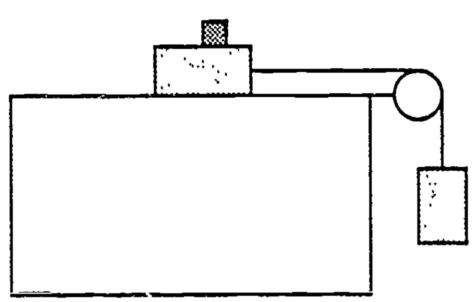


Figure 1. A situation with an "extra" block.

> This one is very interesting because it illustrates a very important point which is that friction isn't always in the wrong direction -- or the right direction. The thing that bothered me right away is the extra mass sitting here [on top]. And the question is: what's going to happen with the extra mass. Well if there's everything accelerating to the right ,then so will this top mass. And on the other hand , your first temptation is to draw a force diagram in your head and you say 'OK the thing's moving to the right, so friction is to the left.' Except that friction is the only force moving it, so right away you reach a problem. The answer is ... that friction opposes relative motion, so the friction's

going to go in whichever direction it needs to point to oppose the motion between the little mass and the big mass it's sitting on. And that happens to be to the right if the big mass is moving to the right.

This protocol illustrates several features of mental models that we will discuss.

First, the protocol indicates that the diagram was parsed into sets of components that operate as systems. The reference to an "extra mass" suggests strongly that the two blocks connected by the string form a constituent unit, and the remaining block is "extra." The process of organizing problem situations into a set of constituent units is one important feature of problem representation.

Second, a process of envisioning apparently aided the subject in constructing a representation. We interpret his comment, "there's everything accelerating to the right," as an inference based on imagining movement caused by gravity pulling the hanging block downward. The comment, "then so will this top mass," did not result from a formal computation or principle – in fact, it opposed a feature of the force diagram that that the subject mentioned. It seems reasonable then, to interpret the subject's inference as an envisionment based on familiar experience such as pulling a wagon that has a box on it.

A third feature is the interplay between representations based on theoretical concepts and representations of objects at the level of ordinary experience. This interplay produced a conflict that the subject resolved by an interesting reformulation of his understanding.

The subject's construction of a force diagram involved an internalized version of a process taught in most physics classes, drawing a free body diagram. In drawing a free body diagram, one draws a force vector for each force acting on a body. The subject said "friction is the only force moving [the top mass];" he apparently focused on lateral forces, since vertical forces are irrelevant to the issue of lateral movement. One common rule of thumb for determining the direction of the friction force vector is "friction opposes motion." Use of that

heuristic could have led to the statement, "OK, the thing's moving to the right, so friction is to the left."

The subject's representation at this point contained a conflict. He had concluded that the lateral force on the top block was friction, pointing to the left. This would lead to an inference that the top block should move to the left – the body should accelerate in the direction of the force. In the informal envisionment, however, the top block was moving to the right, along with the bottom block.

The subject resolved the conflict by revising his mental force diagram. He said, "the answer is -- that friction opposes relative motion, ... the motion between the little mass and the big mass it's sitting on." This reformulation would allow the subject to draw a new force diagram with the force vector for friction in the correct direction, to the right. By revising the theoretical terms of the representation the subject achieved a mental model in which the informal envisionment and the theoretical features were consistent.

2. A General Framework

We propose a framework for characterizing problem representation and reasoning that we call a *relational* framework. The components of the framework are shown in Figure 2. We call this framework the relational framework because it concerns strong, stable relations between different kinds of representation.

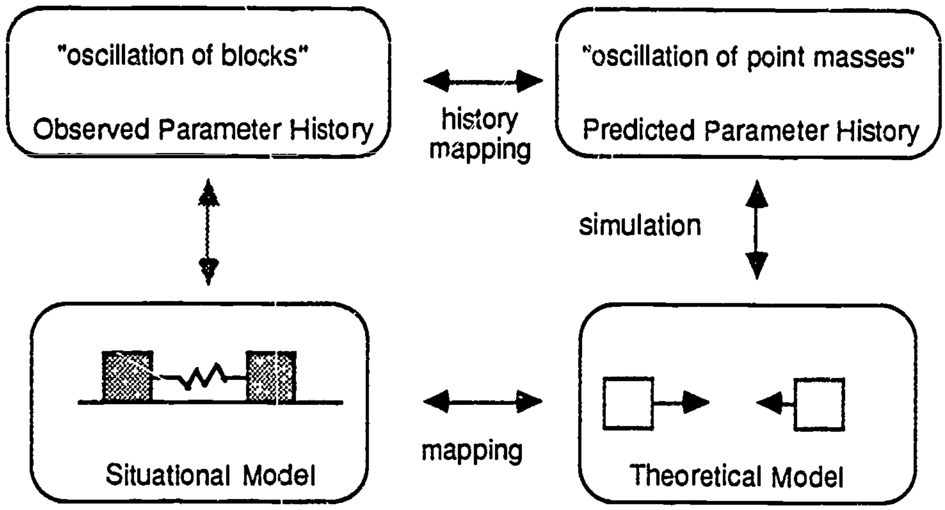


Figure 2: The Relational Framework

A common problem in physics is to provide a coherent explanation of some observable phenomenon. We propose that this explanation can often be characterized as an integrated set of relations among the four components in Figure 2. The components on the left, the situation model and the observed parameter history, correspond to features of the phenomenon that can be observed. The situation model represents objects, properties, and relations in the concrete situation that the physicist considers significant. The parameter history involves features of the dynamic behavior of the system. This is a sequence of the values of some measurable quantities over time. The representation that we call a situation model is like the mental models discussed by Holland, Holyoak, Nisbett, and Thagard (1986) and by Johnson-Laird (1983), and the situation models discussed by vanDijk and Kintsch (1983). The idea of a parameter history follows Forbus (1984).

The physicist's explanation connects the situation model and the parameter history with two additional components: a theoretical model and a set of predictions. The theoretical model, like the situation model, consists of some entities and some relations between them. The entities consist of idealizations such as point masses and forces. Additionally, the theoretical model has properties that make it mathematically tractable, though physicists may also conceive of the model visually. Simulation based on the theoretical model also forms a

parameter history, since the simulation produces predictions about measurable aspects of the dynamic behavior of the model.

The physicist's goal is to establish a stable four-way relationship between the components of the observations and the components of the explanation. We have given names to some components of this relationship in Figure 2. A process of mathematical simulation relates the theoretical model to the theoretical parameter history. This mathematical simulation can be quantitative or qualitative (as in deKleer and Brown (1984)). A mapping of objects and their interrelations links the situation model to the theoretical model. Finally, a mapping of equivalent parameters relates the observed parameter history to the predicted parameter history. When the physicist succeeds, this coherent, interrelated structure is the prize; it is a source of accurate, consistent, and far-reaching predictions and explanations.

We propose that this same relational framework can form the basis of experienced physicists' textbook problem solving. Some subclaims are necessary to make the framework fit. First, note that in textbook problems there are no dynamic events to observe, though the physical situation is described in detail, often with an accompanying sketch. We claim that experienced physicists build a situational mental model based on the textual and graphical presentation of the physical situation, and that they envision this mental model in order to derive a qualitative parameter history. (This process may be compiled for routine situations, in which case the reasoner may use a "canned" parameter history, rather than generating a fresh one.) This qualitative parameter history takes the place of the observed phenomena.

Second, note that in textbook problems, experienced physicists typically use mathematics to derive only a single numerical result from their theoretical model. A single number is not a parameter history. A richer set of predictions from the theoretical model is needed to construct the relational framework described above. We claim that experienced physicists envision mental models of their theoretical representations in order to derive predicted qualitative parameter histories.

3. Further Discussion and Examples

In this section we discuss examples from four protocols that illustrate the relational framework. The subjects were three experienced physicists: a university professor, a graduate student, and a high school teacher. All the subjects had been teaching introductory physics in the past year. The physicists were shown sketches of situation like those found in textbook problems; however the exact configurations were novel to all three subjects. The sketches were not accompanied by any text or additional information. The subjects were not told what the question was. Instead they were asked to explain, "What's happening?"

First, consider the introductory example. According to the interpretation given in Section 1, the physicist recognized a conflict between two envisionings of the motion of the "extra" block. One of his envisionings predicted a parameter history of motion to the right; the other envisioning produced a parameter history of motion to the left. We use the framework of Figure 2. The first parameter history, with motion to the right, derives from a commonsense envisioning of a situational model, involving objects that move according to principles that are familiar from ordinary experience. The second parameter history, with movement of the "extra" block to the left, is based on the physicist's theoretical model, involving the direction of a friction force.

In this example, all the components of the relational framework are in place – both the situational and theoretical models and the derived qualitative parameter histories. The physicist recognized a contradiction implied by the mapping between the situational and theoretical parameter histories. The physicist resolved the contradiction by introducing a refined heuristic for constructing the theoretical model: he expressed this with the statement that "friction opposes relative motion." In this case, commonsense knowledge enabled the physicist to recognize and correct an error in a scientific heuristic.

Now consider a second example based on the sketch in Figure 3.

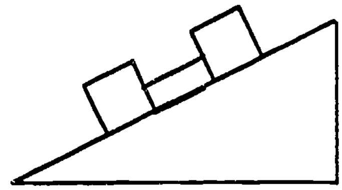


Figure 3. A situation with connected blocks.

> I've never seen this situation before.... We have two masses, I would say with different masses , and let's assume there is a string and not a fixed bar -- with a fixed bar it's of no interest....
>
> Then if there is not so much friction the whole system will go to the left , and if [otherwise] ... no motion at all.... I'm thinking about: if the right block will become higher velocity than the left block?
>
> No, no! That's not possible at all,... they must have the same acceleration because I know that acceleration does not depend on mass.... so that's not an interesting physics problem because they always have the same distance, and it does not depend on mass -- that was my first idea [that it does depend on mass] -- it was a commonsense idea. It does not matter if there is a thread or a bar because both have the same acceleration.

In this protocol, the subject began with some assumptions about the physical situation. This suggests that he was creating a situational mental model. In the second paragraph, the subject envisioned the motion of the objects in the situation under a variety of conditions. The result of this envisioning was a set of situational parameter histories that the physicist was to duplicate in his theoretical parameter history via his theoretical model. Included in the set of parameter histories is one – "the right block will become higher velocity than the left" – about which the subject was not sure.

In the next paragraph, the subject constructed a theoretical mental model of the situation, as indicated by his introduction of theoretical terms like "mass" and "acceleration." The subject recognized a constraint over the possible parameter histories that this model can generate: "they must always have the same distance." This constraint led him to reject the

situational parameter history in which the blocks have different velocities as a commonsense mistake.

The contrast between these two examples reveals an important point: In textbook problem solving, unlike laboratory problem solving, the subject must produce both the observed parameter history and the theoretically-predicted parameter history. Therefore the possibility exists that either the subject's situational knowledge or the subject's theoretical knowledge may be wrong. In the first protocol, the physicist's theoretical representation contained an error, while his situational understanding was basically correct. In the second protocol the opposite was true. This point implies a larger role for commonsense knowledge in scientific reasoning than has typically been described in studies of physics problem solving.

The next two examples illustrate how reasoning strategies propagate along the links in the relational framework. One protocol comes from the diagram in Figure 4:

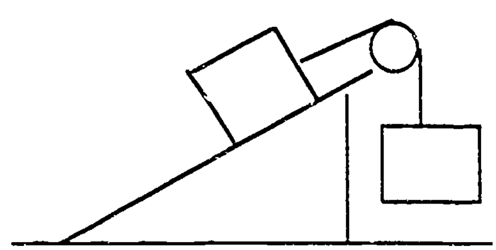


Figure 4. Another situation.

When shown this sketch, the subject immediately began constructing a theoretical representation based on the assumption that the string was always taut. The experimenter noted that the subject had not made that assumption for the sketch in the previous example, though both deal with two blocks, a string and an incline plane. The subject explained his assumption as follows:

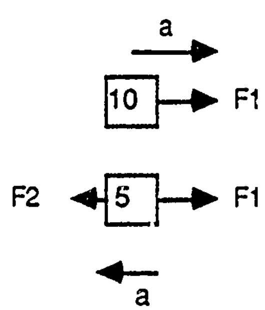
> The string is always going to be taut. I can't tell you why I knew that immediately. I guess because the mass is hanging so that is going to keep the string taut all the time. Unless something weird is going on – I can't imagine what.

The reasoning in this segment notably lacked any theoretical content – no mention of forces or constraint laws. In fact the reasoning was distinctly causal and circular from the Newtonian point of view: the hanging block is not keeping the string taut any more than the string is keeping the block hanging. The physicist did not cite any principles or scientific analyses to back up his assumption. In fact, the physicist appeared to generate confidence in his assumption purely from his inability to imagine the situation being such that the string was not taut. So in this case, we would say that inferences from the situational model propagated directly to the theoretical model without any further checking. We find this is frequently a feature of experienced physicists' reasoning: physicists prefer to make certain inferences, especially those regarding the state of strings, based on situational, rather than theoretical models.

The next example shows a propagation in the opposite direction: from the theoretical model to the situational one. Another physicist discussed the way he viewed sketches like Figure 4 as follows: "I make believe the string's not there and I imagine the whole system as one in which this one [block] is touching that one [the other block] and they're glued together." In this case, the physicist took an operation on the theoretical model, considering rigidly–connected objects to be a single system, and created an operation on his situational model which is equivalent, gluing the blocks together. We suppose that this enables the physicist to make inferences about the system based on commonsense rather than using a model with abstract entities like forces and constraint laws. Through their situational interpretations of theoretical operations, physicists may come to understand the theory in a more intuitively appealing way.

4. Influence of Surface Features

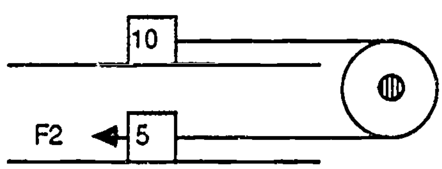
In this section, we report evidence from a small experiment addressing the question: How do surface features affect experienced physicist's performance? We use "surface features" in the sense of Chi, Feltovich, and Glaser's use of the term — the entities and relations visible on the surface of a sketch. We will compare the evidence with predictions from the transitional framework and from the relational framework in a competitive argument.



The accelerations are equal and opposite.
The top mass is 10 kg. The bottom mass is 5 kg.
F2 = 30 N

Find the acceleration.

Figure 5



The pulley, string, and all surfaces are ideal.
The top mass is 10 kg. The bottom mass is 5 kg.
F2 = 30 N

Find the acceleration.

Figure 6

The evidence comes from protocols of three physicists and physics graduate students who were asked to think aloud as they solved two carefully constructed physics problems. Figures 5 and 6 show the problems, which were presented to the subjects sequentially. The problems have identical deep structure; that is, the same mathematical solution applies to both. However, one problem of the pair is presented in free body diagram form with no surface features indicative of a real world situation. The second problem of the pair presents a real world situation, but no free body diagram. Note that this situation, two blocks on parallel tables

connected by a pulley, is not a situation typically found in physics textbooks, though it fits a class of situations often found which can be loosely categorized as "pulley problems."

An alternative hypothesis to the relational framework is that representations are constructed in a linear progression. We call this hypothesis the transitional framework, because it focuses on transitions between representations. (We discuss this alternative more fully in Section 6.) In such a progression, the surface feature representation would be constructed, and then used to construct the free body diagram representation. This framework predicts that once the problem solver develops the free body diagram, the surface feature representation has little or no role in developing the mathematical solution. Both Larkin (1983) and Chi, Feltovich, and Glaser (1981) have made statements to this effect: On this view, surface feature representation plays a diminished role in experienced physicist's problem solving. Thus we would expect equivalent performances from the physicists on the first and second problems, with and without surface features.

The relational framework, on the other hand, emphasizes persistent relations between knowledge at the level of familiar objects and at the level of the theoretical representation. In this framework, surface features form the basis of the situational mental model. This mental model has a continuing role in problem solving after the construction of the free body diagram. Thus we would expect experienced physicist's performance to be impaired on problems of sufficient complexity which omit surface features.

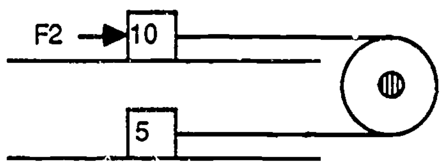
In our observations of three subjects that solved this pair of problems, several differences were noted between the performances on problem 1 and problem 2. First, the solution methods the subjects used for problem 1 differed from those used for problem 2. All subjects solved problem 2 by using a single F=ma equation, combining the two masses for the value of "m", and using the value of the applied force for "F." In contrast, the solutions for problem 1 never directly combined the two masses, but instead used separate, simultaneous F=ma equations for each mass. The solutions for the problem without surface features were

more complex than the solutions for the problem with surface features Second, the subjects required a greater amount of time to solve problem 1, ranging from 2 to 20 minutes, than problem 2, which was solved within a minute by all subjects. Third, the physicists expressed much less confidence in their solution for problem 1 than for problem 2. This was particularly striking in one subject, who after studying the free body diagram for problem 1, declared that he believed the situation to be described inconsistently, and believed solution of the problem, therefore, to be impossible. This subject then set out to prove inconsistency. In the process of writing equations for each body, he made a sign error. Due to this sign error, his equations were in fact inconsistent, supporting his prediction that solution was impossible. However, the subject lacked confidence in this solution, and checked his logic many times. He never found the error.

In summary, on the problem without surface features, subjects used more complex solutions, more time, and had less confidence. These results more closely match the predictions of the relational framework than the transitional framework because the relational framework predicts that surface features continue to be important after the free body diagram is constructed.

In the experiment, the two problems discussed above were followed by two additional problems, designed to clarify a further difference between the transitional and relational frameworks. This difference has to do with the interpretation of surface features. One alternative is to treat surface features fairly literally. For example, the presence of a pulley--grapheme may be enough to trigger a particular solution schema. In the relational framework, on the other hand, surface features signify a real world situation, allowing the construction of a mental model representing the concrete objects in the situation. This mental model can then be run in the mind's eye to generate inferences. In short, the first alternative interprets surface features literally, while the relational framework focuses on the concrete mechanism that the surface features signify.

The two problems discussed above are ambiguous with respect to this difference -- the subjects could be choosing a different solution method for figure 6 because of either the literal surface features or the physical mechanism the surface features signify. The two additional problems given to the subjects help illuminate the difference.

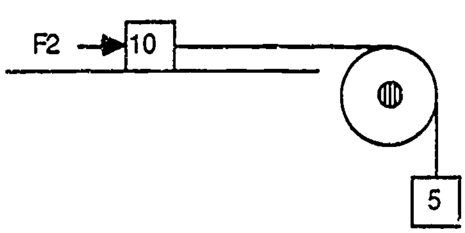


The pulley, string, and all surfaces are ideal.
The top mass is 10 kg. The bottom mass is 5 kg.
F2 = 30 N.

Find the acceleration.

Figure 7

Figure 7 has the same surface features as figure 6, with one difference, the placement of the force arrow. This small difference in surface features results in a very different behavior: the blocks in figure 6 are coupled together by the taut string, but not in figure 7. If surface features are interpreted literally, subjects would be expected to choose the same solution method for figures 6 and 7. On the contrary, subjects immediately noticed the significance of the placement of the force arrow, and thus did not combine the masses of the two bodies in order to solve figure 7 in one equation. This suggests that these physicists used the surface features to build a working mental model of the concrete situation, in addition to any literal uses.



The pulley, string, and all surfaces are ideal.
The top mass is 10 kg. The bottom mass is 5 kg.
F2 = 30 N.

Find the acceleration.

Figure 8

Figure 8 confirmed this result with a slightly different configuration of the same entities. In figure 8, the string can be either loose or taut depending on the applied force. Again the subjects took notice of this difference prior to working out a theoretical solution.

One subject's protocol for figure 8 was particularly interesting. This subject started the problem by noting the two different possible states of the string, taut and loose. He then decided to work on the situation with the taut string first. In doing so, he made a mistake, writing Newton's Second Law as "a = m/F" rather than "a = F/m." After proceeding to a numerical solution, he reasoned that if the applied force were great enough the acceleration should approach free fall, which would be the boundary case between the taut and loose states. Next he consulted his equations, expecting to find the acceleration increase with the applied force. However, he noted that the force was in the denominator, which would indicate decreasing acceleration with increasing force. This disparity between his commonsense expectations and his theoretical model led him to find and correct his error. Without enough surface features to allow the subject to make inference about the behavior of the concrete mechanism, it is doubtful that he would have corrected his theoretical model.

5. Discussion: Constructing a Relational Representation

In this paper, we have introduced the relational framework, a framework for looking at physics problem solving. This view identifies strong, stable relations between representations of concrete, familiar situations and representations of abstract theoretical entities. In this view, interactions between concrete and abstract representations, as well as between commonsense and scientific knowledge are central to physicist's reasoning about novel physics problems. We have presented evidence from studies of experienced physicists that lead us to this view. In this section, we discuss characteristics of cognitive processes that are required in the relational framework and the kinds of mental models that we propose as the representations of physics problems by experienced physicists.

5.1. Qualitative Process Models. Mental models have been discussed extensively elsewhere (Gentner and Stevens, 1983; Holland, et. al., 1986; Johnson-Laird, 1983). Basically, mental models are internal representations that duplicate properties and interrelations of external reality at the level of individual objects. Smith (1987) discusses one property of mental models that he calls "absorption." In representations that have the property of absorption, symbolic objects are like the objects they denote. Rules for combining and transforming symbolic objects and relations in the representation have results that are like the results of transformations of real objects and relations.

The mental models observed in physics problem solving have particular characteristics; they are qualitative, contain both physical and abstract objects, and represent dynamic processes. Representations of this kind have been studied by AI researchers in their investigations of qualitative reasoning (e.g., Forbus, 1984; deKleer & Brown, 1981; 1984). Our synthesis of these works follows:

A qualitative process model is a structure made up of connected qualitative process pieces (QPP's). A QPP describes a single process. In order to model a complex event, QPP's may be connected in serial (when the state of the system changes character at some point in time), in parallel (when two semi-independent systems are active), or embedded in one another (for more detailed views of subprocesses). Because they encode causal relations among parameters, QPP's facilitate qualitative reasoning. Abstracting the features of Forbus's (1984) and deKleer and Brown's theories (1984) yields an elementary QPP with the following components:

- some objects and their relevant parameters
- some parameter conditions which describe when the QPP is valid
- some qualitative causal relations, which describe how changes in parameters propagate
- some constraining relations, which reduce the degrees of freedom in the envisioning process

As Forbus has shown, a knowledge structure with these characteristics can represent both commonsense and scientific process models. Thus we take this structure to be the base representation of the situational model and the theoretical model. From qualitative process models of the situation, a reasoner could envision observed parameter histories, while qualitative process models of theoretical concepts could lead to predicted parameter histories.

We conjecture that much of knowledge in physics consists of knowledge of individual QPP's and of the connections between them. In addition, knowledge includes general principles and facts with applicability in many different processes. By encoding knowledge in the form of mental model pieces rather than individual propositions, physicists gain the coherence of a larger, more redundant knowledge structure. Moreover, increased efficiency results because physicists can now parse a problem situation according to mechanisms rather than individual objects.

For example, in the protocol of Section 1, the subject apparently parsed the situation into two systems with separate mechanisms. The first, we could call the Atwood's Machine QPP. An Atwood's machine is a pair of blocks suspended on the opposite sides of a pulley via a rope. Based on qualitative reasoning about this process, the subject could conclude that the system of the two blocks, string, and pulley would accelerate to the right and down. The second mechanism would be static friction, which operates between the block on the table and the top block. Static friction introduces a sideways force on the top block which causes that block's motion. These two processes are connected in parallel, because they are active simultaneously. Moreover, they are connected by a shared parameter, the force of friction, which acts on both systems.

In addition to the parallel connection between QPP's evident in this protocol, we might expect several transitional and embedding connections. For example, we can expect physicists to know that if the friction between the block and the table were great enough, no motion would occur. Thus a change in the friction coefficient between the block and table

could result in the process changing from one of static equilibrium to one of accelerated motion. Furthermore, a physicist might have embedded QPP's which represent the Atwood's Machine Process at varying levels of detail. For example, one physicist in our study explained that he could consider any "two-objects-connected problem" as either a single rigid body (in which case no tensions enter the analysis) or as two bodies connected by the tension in a string. In this case, the rigid body analysis could be connected to the string tension analysis by an embedding link, signifying that the latter process was a more detailed view of the former. Such links enable the reasoner to recall processes that are related the current representation in order to make inferences about different conditions or at different levels of detail.

5.2. Qualitative Reasoning Processes. Forbus (1984) and deKleer and Brown (1984) both give examples of algorithms that build qualitative process models for given situations. We hypothesize that physicists build the representations of the kind described in the relational framework using some of the component processes in these algorithms, without necessarily following a set flowchart. Instead subjects may build a relational framework via iterative refinement of the individual representational structures. The reasoning processes subjects use in iterative refinement may be captured in rules of several types.

For example, subjects may use envisioning rules to produce a predicted parameter history from a model. Envisioning rules may include very general rules for propagation of qualitative changes, as in Qualitative Process Theory. In addition, domain specific rules are likely to be present — "if equal and opposite forces act on a body it will not accelerate" is an example. Remember that subjects may need to generate parameter histories for the situation model as well as the theoretical model. In that case, envisioning rules may work from episodic memory of real world events. However, commonsense general principles, for example "heavy things fall faster," may figure in envisioning the situation model as well.

Another class of rules, comparator rules, are necessary to provide feedback about the consistency of the relational representation. Particularly relevant are comparisons between

the structure of the situation model and the theoretical model, and between the observed and predicted parameter histories. In addition to envisioning and comparator rules, two other classes of rules are necessary: selector and finder rules.

Selector rules retrieve QPP's and facts from long term memory to serve as the basis of the situational and theoretical models. The triggering conditions of these rules can vary widely. As Chi, Feltovich, and Glaser (1981) proposed, literal words and objects present in the surface feature sketch may cue knowledge retrieval. For example, the presence of blocks, a pulley, and a string might cue an Atwood's Machine QPP. Dynamic features might also serve as cues. For example, motion around a circle might cue an Circular Motion QPP. In addition, mathematical features might serve as cues. For example, one experienced physicist we observed used a spring process model in his analysis of an abstract situation, apparently because the given equation resembled the form of the equation for a spring. Finally, characteristics of the mechanism signified by the surface feature sketch may cue processes. For example, an Atwood's machine can be built with a rope, a wire, a chain. While these systems are different at the surface level, at a deeper level they all fit into an equivalence class defined by their mechanism of operation.

Once a set of QPP's is retrieved from LTM, finder rules serve to instantiate them. Finders, too, may draw information from a wide variety of sources. A simple finder might instantiate a model parameter with a literal piece of given information (such as "the initial velocity is 5 m/s"). Other finders might use general facts. For example, in one of the protocols discussed above, the physicist apparently inferred that the acceleration of both blocks on the incline plane would be the same from a general fact about blocks on incline planes. Another class of finders might use procedures to generate the needed information. Some of these procedures follow general principles. For example, most students learn a procedure for adding up the forces acting on a block on an incline plane to determine the acceleration. This procedure, while based on a theoretical model, may be compiled into a form which does not

require building a detailed representation of the acting forces. More complex finders might use the structure mapping between models in relational framework to map information from representation to representation. For instance, the reasoner might know that a block sitting on a table will not move in the vertical plane. This information might be used to instantiate the vertical acceleration as having the value zero.

6. General Implications

We conclude with some comments about relations of the scheme that we have presented to other recent discussions of representations and mental models in problem solving.

First, the relational framework for a textbook problem, including reciprocal interactions between situational and theoretical mental models with associated qualitative parameter histories, contrasts with an earlier model developed by McDermott and Larkin (1978). Their analysis involved a four-stage framework for understanding physics problem solving. Each of McDermott and Larkin's stages captures a level in a progression of representations: a statement of the problem in words, a sketch of the situation, a sketch of abstract entities (e.g. forces), and finally mathematical representation. In this framework, problem solving knowledge has two components: construction rules, which build the representation at the next stage from the representation at at the previous stage, and elaboration rules, which enhance the information within a stage. We label this framework the *transitional* framework, because it describes a transition from representation to representation. Figure 9 shows the four stages with sketches corresponding to the representations that are formed for a problem involving a spring.

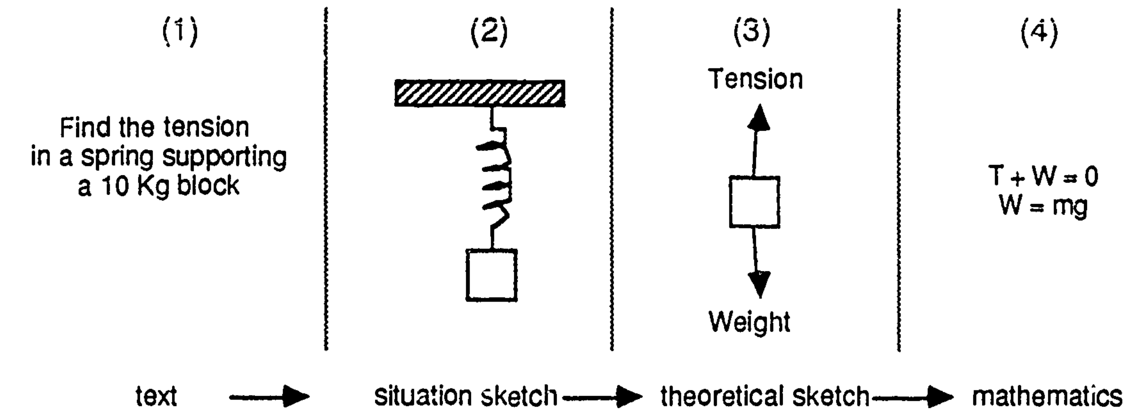


Figure 9: Four stages of the transitional framework

We believe that the relational framework is better adapted than the transitional framework for explaining several phenomena of physics problem solving including:

- envisioning
- problem and representation re-formulation
- ability to adapt to novel situations
- ability to recover from errors
- the importance of understanding the real-world situation

While we have argued that a relational framework best explains these phenomena, we believe that a transitional framework best explains other phenomena. We do not favor one framework over the other, but recognize the existence of both intertwined in physicists' problem solving behavior. Figure 10 shows how these frameworks might fit together.

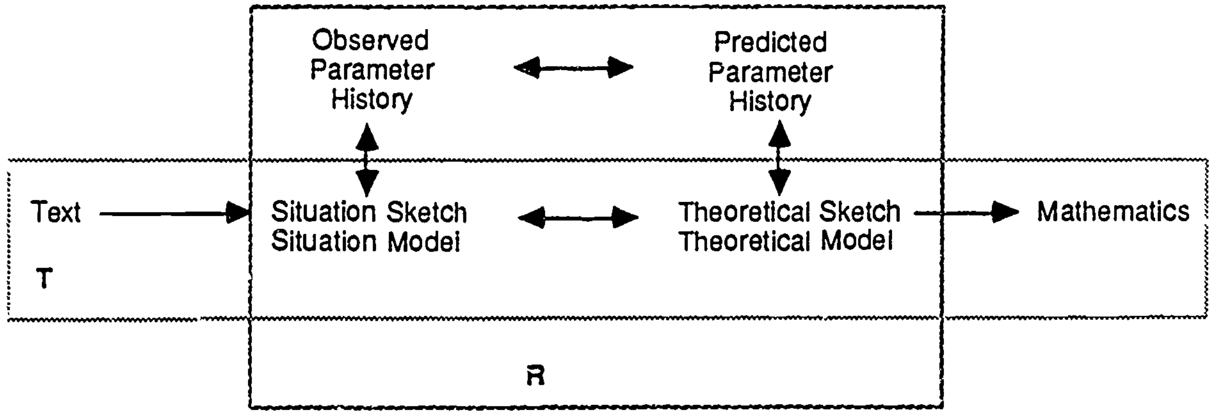


Figure 10: Transitional and Relational frameworks combined

Our analysis is also related to recent discussions of mental models, notably by Johnson-Laird (1983) and by Holland et al. (1986). Johnson-Laird analyzed syllogistic reasoning, involving tasks such as saying what follows from premises such as "Some actors are bakers, and no bakers are carpenters." His major conclusion was that individuals construct models of the premises, including tokens that correspond to individuals with properties that the premises refer to. There might be two tokens that are actors and bakers, two more tokens that are actors but not bakers. One of the actor-tokens could be a carpenter and another token could be added that is a carpenter, and neither an actor nor a baker. This would lead to the following tableau:

```
a -- b
a -- b
a
a ------ c
         c
```

Based on this mental model, an individual might conclude that some actors are not carpenters, or that some actors are carpenters, or that some carpenters are not actors. Only the first of these is a valid conclusion from the premises. To eliminate the others, the reasoner

must construct additional models that differ from the one shown above, but that are also consistent with the premises. Johnson-Laird attributed errors in syllogistic reasoning to limitations of the process of generating models and holding them in memory.

In another recent discussion of mental models, Holland et al. (1986) proposed a hypothesis of a default hierarchy of rule-based models. A situation is categorized as specifically as possible, and situation-specific rules are used to interpret the situation and make inferences. If the specific model does not produce an adequate answer or interpretation, more general categories and rules are applied.

One example discussed by Holland et al. (1986) is an analysis of logical reasoning by Cheng and Holyoak (1985) and by Cheng, Holyoak, Nisbett, and Oliver (1986). These studies focused on a task introduced by Wason (1966) in which participants choose items to test a conditional proposition. For a statement "If A then B," tests should include both examination of A's, which must be B's, as well as non-B's, which must be non-A's. Cheng and her colleagues found that performance was facilitated when a pragmatic reasoning schema such as permission could be applied. For example, performance on a task based on "If the envelope is sealed, there must be a 2d stamp" was better when there was an explanation that sealing the envelope was allowed for first-class mail, which required a 2d stamp, than when the rule was given without an explanation.

Holland et al. (1986) contrasted reasoning in tasks involving logical propositions and tasks involving statistical inference. They characterize the rules of statistics as graceful, in that they build on and support rules that individuals know, and the rules of logic as alien, in that they are inconsistent with known rules. Formal training in inferential rules using the conditional proposition was ineffective unless it was accompanied by training using examples (Cheng et al., 1986), but in the case of the principles involved in the law of large numbers, both formal training and training involving examples were effective, and their combination was more effective than either kind of training by itself (Fong, Krantz & Nisbett, 1986).

The relational framework provides an alternative in which formal rules of inference and empirical rules based on experience interact productively. This extends the ideas of Holland et al. (1986), in which abstract rules are used when more specific rules fail to give adequate answers. In the relational framework that we have discussed, abstract rules supply constraints that are used in construction of the representations of specific situations and problems.

While our relational framework adds to the scheme of a default hierarchy, it is consistent with the main ideas of Holland et al.'s (1986) analysis. In particular, the idea of pragmatic schemata provides a plausible hypothesis about the process of constructing situation models. Causal relations, and schemata of physical mechanisms, provide bases for organizing the features of problems, enabling the use of inferential rules that the problem solver can then apply.

The relational framework has implications both for education and for future research. Physics courses typically emphasize quantitative rather qualitative solutions, and formulas rather than models. In addition, rather than helping students integrate their commonsense understanding of a situation with its theoretical model, physics instruction often implores students to leave the commonsense and familiar behind. If the relational framework is as important to expert-level problem solving as our studies indicate, this emphasis on the quantitative, formulaic, and theoretical prevents students from attaining competence. One of us, Roschelle, is currently exploring the use of computer-based simulations to help students acquire qualitative mental models of physics concepts. (See Roschelle 1986.)

In terms of basic research on cognition, we look for a better understanding of the interaction between situational and theoretical mental models in physics problem representation to be important in several areas. As discussed above, this interaction provides critical feedback for error-correction both in the situation model and the theoretical model. In

addition, case analysis, the strategy of considering several possible cases for problem situation in parallel, may have its basis in the understanding of the different states of behavior for a mechanism that qualitative analysis of both the situational and theoretical model provides. Finally, this interaction appears to be essential to physicist's planning for novel problems.

Acknowledgements

Thanks to Sue Allen, John Seely Brown, Jill Larkin, Peter Pirolli, Alan Schoenfeld and Lisa Quinn for comments and suggestions. Special thanks to Andy diSessa for insight and enthusiasm.

References

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology, 17,* 391-416.

Cheng, P. W., Holyoak, K. J., Nisbett, R. E., & Oliver, L. M. (1986). Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology, 18,* 293-328.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 87-119.

deKleer, J., & Brown, J. S. (1981). Mental models of physical mechanisms and their acquisition. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 285-310). Hillsdale NJ: Lawrence Erlbaum Associates.

deKleer, J., & Brown, J. S. (1984). A qualitative physics based on confluences. *Artificial Intelligence, 24,* 7-84.

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18,* 253-292.

Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence, 24,* 85-168.

Gentner, D., & Stevens, A. (Eds.), *Mental models*. Hillsdale NJ: Lawrence Erlbaum Associates.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning, and discovery.* Cambridge MA: MIT Press.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge MA: Harvard University Press.

Larkin, J. H. (1983). Expert representations of physics problems. In D. Gentner & A. Stevens (Eds.), *Mental models.* Hillsdale NJ: Lawrence Erlbaum Associates.

McDermott, J., & Larkin, J. H. (1978). Representing textbook physics problems. *Proceedings of the second national Conference, Canadian Society for Computational Studies of Intelligence.* Toronto: University of Toronto.

Roschelle, J., 1986, *The Envisioning Machine: Facilitating Students' Reconceptualization of Motion,* Xerox Palo Alto Research Center Working Paper.

Smith, B. C. (1987). The correspondence continuum. Stanford University, Center for the Study of Language and Information.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology.* Harmondsworth: Penguin.